Summary Analytics

www.smr.ai
415-234-3553
info@smr.ai

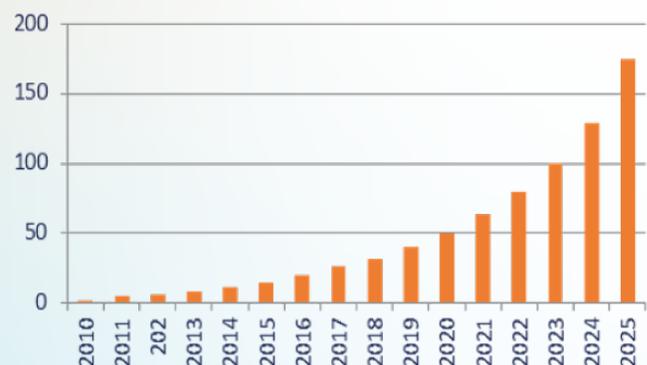## Summary Analytics for Healthcare

An October 2019 survey of more than 900 health-care professionals by MIT Technology Review Insights found that 75% of medical staff who have artificial intelligence (AI) agree that it has enabled better predictions in the treatment of disease, 78% report that their AI deployments have already created workflow improvements, and 80% believe AI is helping them improve revenue opportunities.[1] According to Gartner, AI augmenting human decision-making will create US$2.9-trillion of business value and 6.2-billion hours of worker productivity globally in 2021.[2] This research and forecast quantify the high expectations for AI and create urgency for healthcare management and their supporting IT teams to implement AI today. It's not just about the risk of falling dramatically behind the competition – it's about saving lives.

Yet, Dimensional Research found that 96% of companies have run into training related problems – including data quality, labeling required to train an AI system, and building model confidence – with 78% of their machine learning (ML)/AI projects stalling at some point before deployment.[3] Pactera Technologies's survey showed that 85% of AI projects ultimately fail.[4] In fact, it was widely reported how the MD Anderson Cancer Center at University of Texas wrote-off $62-million on a failed AI project with IBM.[5] How do you get the benefits of AI when the odds are against you with patent lives and large investments at stake?

Developing an effective AI model requires extensive repetition with trial and error analysis of historical data. But often the historical datasets are overwhelming in size and need manual labeling before the models can be tested. For many healthcare applications, this data labeling requires extremely expensive and highly skilled specialists, such as oncologists or radiologists. With Summary Analytics's mathematically proven artificial intelligence techniques, you can shrink the datasets through summarizing and prioritizing without loss of fidelity – delivering better insight while reducing time and cost, and significantly reducing the amount of manual data labeling required. This minimizes the common problem of operator fatigue errors in data labeling and the resultant errors in the models, while also reducing cost and time to completion.

Annual Size of the Global Datasphere (Zetabytes)

Source: Data Age 2025, from IDC Global DataSphere, Nov. 2018

Likewise, Summary Analytics can help with training your AI models. The computational power required to train state-of-the-art AI models is doubling every 3.4 months[6] as Moore's Law continues losing steam, no longer doubling processor performance every 18-months. So far, this problem has been addressed with machine learning algorithmic advances and increased parallel compute power. These help, but more is needed to stop runaway AI analytics costs and delays. A new complementary tool is needed, adding "informational efficiency" to the process. That tool is Summary Analytics. Our software-as-a-service (SaaS) offering summarizes and prioritizes data sets before running expensive analytics. Summary Analytics enables early model testing on significantly reduced and prioritized datasets, while saving larger (but still reduced) datasets to be used for final optimization of the model. Summary

Summary Analytics for Healthcare

Analytics eliminates redundancies in your data. More than just deduplication, we do this even among massive numbers of unique records. We eliminate the unnecessary and shrink the haystack so finding the needle of insight is faster and less expensive. Of course your data is more complex and dynamic than a single needle in a haystack, with new streams and data churn constantly adding new and removing old hay and needles. Worried about what to do with old data? We help create a data hierarchy to focus on the important data, whether old or new.

How does it work? Professor Jeff Bilmes from University of Washington in Seattle developed proprietary calibrated submodular (CaSM) functions which mathematically analyze and order data along the lines of diminishing marginal returns. We automatically prioritize the data in terms of its biggest contribution to the information content of the entire data set, and then relegate redundant data to the end. CaSM functions are extremely processor efficient – orders of magnitude faster than typical AI algorithms. They don't replace AI algorithms, our CaSM functions just make machine learning run much faster since the data sets are vastly smaller but still contain all the important information.

In no other industry do ML and AI have the impact that they do in healthcare – from drug discovery and genetic research, to assisted clinical diagnosis and treatment, to hospital operations and insurance. These technologies are enabling breakthrough discoveries, reduced errors, worker productivity and most importantly, improved patient outcomes. But developing the ever-increasingly complex AI models and training these models is getting more expensive as the datasets required to do it right grow in size. Summary Analytics works great on electronic health records, but that's not all. Whether biological signals, genetic data, sensor data, or even radiological images or video streams – the bigger or more redundant the data, the more Summary Analytics can reduce time and costs and dramatically improve your odds of AI success for your patients.
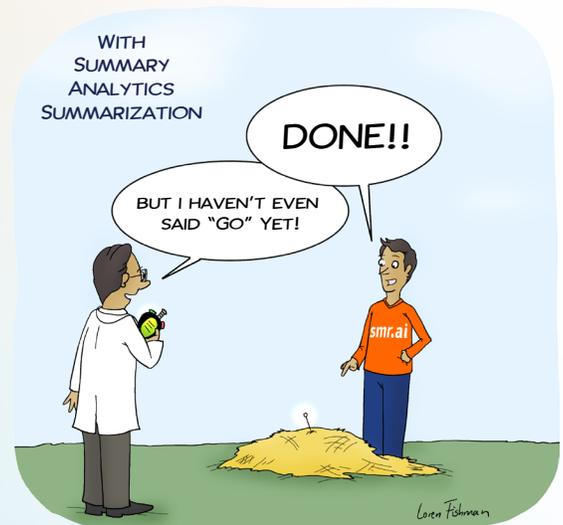
Bigger data? Bring it on!

[1] https://www.technologyreview.com/hub/ai-effect/?utm_source=GS_Download &utm_medium=newsletter&utm_campaign=us_ge_healthcare_reportreinvented&utm_term=US_GE_Healthcare_ReportReinvented_Newsletter_Download_UntangleComplexity_1&utm_content=Photo_Link_Ad#ge-ai--sec1

[2] https://www.gartner.com/en/newsroom/press-releases/2019-08-05-gartner-says-ai-augmentation-will-create-2point9-trillion-of-business-value-in-2021

[3] https://content.alegion.com/dimensional-researchs-survey

[4] https://www.techrepublic.com/article/why-85-of-ai-projects-fail

[5] https://spectrum.ieee.org/biomedical/diagnostics/how-ibm-watson-overpromised-and-underdelivered-on-ai-health-care

[6] https://www.technologyreview.com/s/614700/the-computing-power-needed-to-train-ai-is-now-rising-seven-times-faster-than-ever-before/?utm_source= newsletters&utm_medium=email&utm_campaign=+the_download.unpaid.engagement